IIMC Library Clipping Service | Source: News: Online: Forbes India | Dated: 22-06-2021

In defense of a transparent economy for data capitalism.

BY RANJAN PAL, JON CROWCROFT, MINGYAN LIU, SWADES DE, BODHIBRATA NAG

The value of data to modern businesses helps them to know customers better, make better decisions, and boosts ROI, and form mutually beneficial and profitable business relationships with other organizations.

We are in a digital economy where data acting as the new oil is more (fundamentally) valuable than ever, and the modern key to smooth and efficient functionality of everything from the government to local companies. Today, an unprecedented amount of analysable information on humans, things, and nature opens up vast opportunities for accelerated insights, innovation, economic growth, and advances in scientific and medical research.

The value of data to modern businesses helps them to know customers better, make better decisions, boosts return on investment (ROI), and form mutually beneficial and profitable business relationships with other organisations. According to a United Nations Financial Trade Quarterly (FTQ) report of 2019, the five largest data firms in the world today: Apple, Amazon, Facebook, Google and Microsoft—are actors in the digital data economy (compared to only one of them being an actor a decade ago) with a combined market value of nearly $4 trillion. This amount represents approximately 20 percent of market capitalisation in the US, and more than the annual gross domestic product (GDP) of India.

Concerns with the existing form of economy

Interestingly enough, the people whose raw data is driving the fourth industrial revolution play a rather passive role in the modern digital economy as they are often left out of the value chain that transforms their raw data into huge monetary benefits.

In addition, this digital economy brings an added disadvantage to common people in the form of privacy risks. As they say, if this is the age of information, then privacy is the issue of our times.

Activities considered private in the past or shared with a few, now leave snippets of behavioral data that expose our privacy-sensitive traits, beliefs, and intentions. We communicate using e-mails, SMSs, and social media; find partners online; read and write on the cloud; seek response to sensitive questions using search engines; use geotracking guides to navigate on the road; and celebrate and mourn on social media. Through these and numerous other online activities, we reveal information—both knowingly and unintentionally—to one another, to commercial entities, and to our governments.

The monitoring of personal information is ubiquitous, and its storage durable enough to render one's past undeletable. Accompanying this acceleration in data collection are steady AI (artificial intelligence) advancements in the ability to aggregate, analyse, and draw sensitive inferences from an individual's data [eg via (non)federated learning].

As a glaring example of a major privacy fallout resulting from these advancements that are catalysed by the demands of the economy, the Cambridge Analytica scandal and the influence of the 2016 US elections demonstrated that individuals are increasingly at a privacy risk and likely to be manipulated through big data aggregating and analysing firms. More generally, as valuable human experiences-bearing personal data account for a large part of big data in the mobile and Internet of Things age, they subsequently give rise to a new form of exploitative capitalism ie "surveillance capitalism", in which raw and free information on society individuals is systematically and often shamelessly, and unfairly analysed using powerful AI tools to sell predictions on their behaviors to targeted advertising agencies, much like behavioral futures contracts.

An Alternative Transparent and Potentially Efficient Data Economy

We argue in favor of an alternative transparent data economy that has the potential to be economically efficient at best, in which people will be suitably paid (and explicitly informed) whenever their data will be used for revenue-generating products and services. The basis of our argument are pivoted on the following viewpoints:

1. Paying for data puts economic pressure on online services (that currently incur virtually zero data collection and processing costs) to apply data minimisation (DM) principles, i.e., to collect and process only the minimum amount of data necessary for their operation, that are mentioned in the General Data Protection Regulation (GDPR), thereby mitigating privacy risks.

2. Currently, no existing data protection regulation (DPR) in the world establishes a property right over personal data or confers that right to data subjects. Although the GDPR seems to be inspired by some property-like rights logic, as is evidenced by its introduction of the principles of data portability and the right to be forgotten, it stops short of recognising property rights over data.

The rule is that data can be possessed by the entity collecting it without any property right being affected. A property right would be the first step towards a transparent and fair data economy, and would involve providing the data subject with the use of, as well as the possibility to sell, their personal data, license it to someone for profit and/or use their data as security/collateral for raising capital, as is the case with intellectual property rights, and also be a catalyst to DM principles being adopted.

Although some may consider data to be an intangible asset that may be protected by property rights, currently this is not possible with personal data. It is the exploitation of this personal data acting as a public economic good that certainly creates value, however, this is entirely, or rather overwhelmingly and unfairly, captured by the entities that

harvest this data—one such entity being digital platforms. Property rights would ensure appropriate ownership of data, necessary to realising complete contracts between buyers and sellers in an efficient data economy.

3. As personal data is a public good exploited by digital platforms and their likes, negative impact (termed as 'negative externalities' in economics jargon) is generated on the consumers due to the non-transparency of its use (via incomplete data contracts at best) in a highly complex commercial network. In the presence of DPRs and property rights, paying people for their data will contribute to the (universal) guaranteed minimum income principle in neo-classical economics and be an alternative to labor-based compensation in the future in which most work will be done by machines—a survey conducted recently estimated that if fair remuneration algorithms that take into account both economic and social costs1 are set in place, (a) a family of four in the developed American economy could earn upto $20,000 per year from their data—clearly indicating the positive impact such remuneration schemes would have on developing and under-developed economies. We emphasise here that the remuneration value is reflective of the amount that enables cancellation (termed as 'internalization' in the economics jargon) of negative economic externalities due to sale of personal data to sustain an economically efficient data economy—something that can be only be extracted with property rights-driven and increasingly complete economic contracts.

4. Business models and machine learning algorithms have zero value without human data, and thus should be taxed for collecting the latter—a viewpoint shared by industry leaders such as Bill Gates, Mark Zuckerberg, and Elon Musk.

5. Online services market is not a zero-sum game—increasing the profits of people by paying for their data does not have to harm the profits of online services, and might actually result in higher quality and quantity of data being shared by individuals in many application scenarios and lead to a boost in investment and innovation in technology. As an example, instead of just collecting product pages that users visit on an e-commerce site, the users that decide to opt-in to receive remuneration for their data may further release the amount of time spent at each page as well as the local interaction patterns, that may not be visible through the standard cookies and other mechanisms used today. Moreover, an efficient data economy would improve transparency of price and its impact on customer (e.g. advertisers) business (via improved traded-data quality)—thereby preventing users and intermediaries alike from selling fake user stats. With the current broken market, you get a much less precise signal, and run an increased risk to be sold fake behavior.

Bottomline, the idea of paying people for their data, and with their permission may be a step ahead of our times; still our arguments in favour of alternative transparent data economies are rising fast as they show enough promise to benefit society as a whole—especially, when in a recent effort it has been established via experiments that the GDPR may not be leading to a data minimisation principle. On this latter point, we run the risk of running into a privacy/trust-centric tragedy of the commons problem on the web if we fail to properly apply data protection laws such as the GDPR to monitor and detect unwanted violations of privacy that may arise due to greedy 'all-you-can-eat' practices by online services. To this end, a transparent data economy can play a key role in complementing GDPR-like policies to prevent 'parasitic' online services from gathering anything and everything without going out of business, thereby benefiting society.

However, a challenge worth mentioning with respect to such an economy is the fair and heterogeneous valuation of personal data over time and space. A possible solution is to embrace time-dependent bundled pricing of multi-dimensional data based on the statistical distribution of their use in time and space—ideas and formal models for which have been recently proposed.

A recent work formalised a provably-optimal, distributed, and conditionally efficient micro-economic market mechanism (validated via a small pilot experiment) to realise a version of such a transparent data economy—an important necessary condition being the assumption of the enforcement of complete data contracts requiring property right laws as mentioned earlier. Specifically, transparent trade is modeled through behavior-driven voluntary compromise in privacy by data sellers, in return for rewards, post being explicitly buyer-informed about data use.

The Privacy Paradox Favoring a Transparent Data Economy

Many would agree that imposing a data economy is outright creepy from a privacy viewpoint. However, human behavior sometimes suggests otherwise even without the presence of a regulated economy. The well known privacy paradox introduced by Susan Barnes in 2006, and advocated for many scenarios ever since, have clearly shown a discrepancy between online privacy concerns and privacy behaviors, i.e., even when users/individuals have substantial concerns with regard to their online privacy, they engage in self-disclosing behaviors that do not adequately reflect their concerns. We em- phasise here that in the modern smartphone age, these behaviors are often supplemented by psychology-driven actions (eg, binary opt-in policies set by apps, individuals preferring free apps over paid apps, a significant population of users being neutral to cookie downloads) by individuals and data extractors alike that increase the privacy risk for online individuals. In addition, field experiments conducted in the late 1990s and early 2000s—the pre smartphone age, have established the fact that direct monetary incentives do bias (privacy-sensitive) people to sell their data, thereby adding more weight to (a) the privacy paradox, and also (b) the willingness to embrace a transparent data economy as described earlier.

At the same time, on the contrary, there have been experimental studies not showing enough evidence in favor of the privacy paradox, and conclude that the perceived privacy risk—a construct that loosely resembles privacy concerns—was associated significantly with the respondents' amount of self-disclosure.

To this end, two recent contributing technical factors that rationalise the non-advocacy of the privacy paradox are (a) the rise of privacy-enhancing browsers (PEBs) such as Duck-Duck-Go, Vivaldi, and Brave, and (b) the increased use ad-blocking technology as a means to 'push-back', alongside deciding to gain utility from apps— both of which provide a choice to privacy-sensitive end-users in managing their privacy concerns upfront.

One could argue that there is significantly smaller percentage of PEBs in use when compared to non-PEBs (e.g., Google), mainly due to 'winner-takes-all' economics; however, that doesn't mean that an efficient market for users' data wouldn't work. The status-quo is evidence of failure to regulate misbehaviour by non-PEB oligopolies. If regulation can be effective, we could have dense markets for both non-PEBs and PEBs.

However, despite the rise of ad-blocking technology, ad blocking firms like Eyeo, maker of the popular AdBlock Plus product, has achieved such a position of leverage that it gets Google et.al., to pay it to have their ads whitelisted by default, under its self-styled 'acceptable ads' program which clearly goes against the core functionality principle of ad-blockers, and, in a sense, nullifying the advantage of users not wanting to fall prey to the privacy paradox.

In addition, very recently, researchers have shown via experiments conducted on mobile users that paying for apps (initially hypothesised to mitigate mobile user privacy risks) does not guarantee better privacy either. On the human behavioral side, one could draw connections between diverse streams of empirical research on privacy behavior to showcase why privacy risks are inevitable. We use three themes to connect insights from social and behavioral sciences: (i) peoples' uncertainty about the consequences of privacy-related behaviors and their own preferences

over those consequences, (ii) the context-dependence of people's concern, or lack thereof, about privacy, and (iii) the degree to which privacy concerns are malleable—manipulable by commercial and governmental insights. On (i), people are often uncertain about what information other people, firms, and governments have about them and hence, they are uncertain about how much information to share and end up sharing more than necessary at times.

On (ii), context-dependence means that individuals can, depending on the situation, exhibit anything ranging from extreme concern to apathy about privacy, i.e., the way humans construe and negotiate public and private spheres is context-dependent because the boundaries between the two are murky.

Finally, on (iii), with the emergence of the information age, growing institutional and economic interests have developed around disclosure of personal information and some entities have developed expertise in exploiting behavioral and psychological processes to promote disclosure, and such efforts (e.g., mentioned above in the context of mobile apps) play on the malleability of privacy preferences.

Thus, in the wake of inevitable privacy risks, the idea of a transparent data economy can promote economic parity between raw data owning individuals and its buyers, provided there is a scalable security and privacy enhancing technology atop which it can be implemented.

Ranjan Pal, EECS, University of Michigan,Ann Arbor, USA.

Jon Crowcroft, Computer Laboratory, University of Cambridge, UK.
Mingyan Liu, EECS, University of Michigan, Ann Arbor, USA.
Swades De, Electrical Engineering, Indian Institute of Technology, Delhi, India.
Bodhibrata Nag, Operations Management, Indian Institute of Management, Calcutta, India.

Source: https://www.forbesindia.com/article/iim-calcutta/in-defense-of-a-transparent-economy-for-data-capitalism/68691/1